REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)

2. REPORT DATE -

3. REPORT TYPE AND DATES COVERED
FINAL 25 Jun 97 - 30 Jun 00

4. TITLE AND SUBTITLE
Hardware Realization of a Ethernet Packet Analyzer Search Engine

5. FUNDING NUMBERS
DAAG55-97-1-0339

6. AUTHOR(S)
Lionel L. Ramos

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Texas - San Antonio

8. PERFORMING ORGANIZATION
REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U. S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING
AGENCY REPORT NUMBER
ARO 36888.1-CI-AAS

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This project encompasses the design and implementation of a hardware realization of Ethernet packet analyzer with a customized search engine that is specific for the home automation industry. This analyzer will be at the gateway of a network and analyze Ethernet packets as they go by. It will keep a record of all packets sent to the gateway. I will also perform a real-time specialized search of the data packets for information that is for the home automation and not the computer network. This system is a stand-alone real-time network analyzer capable of decoding Ethernet protocols. The research project includes reviewing several different types of control architecture, clarifying search engine, protocols, and network interface controllers.

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

UNIVERSITY OF TEXAS AT SAN ANTONIO
COLLEGE OF SCIENCES AND ENGINEERING
DIVISION OF ENGINEERING

ASSERT RESEARCH REPORT

"HARDWARE REALIZATION OF A ETHERNET PACKET ANALYZER SEARCH ENGINE"

By
Lionel L. Ramos
Graduate Student
Fall 2000

Augmentation Awards for Science and Engineering Research Training (ASSERT) RESEARCH REPORT

"HARDWARE REALIZATION OF A ETHERNET PACKET ANALYZER SEARCH ENGINE"

INTRODUCTION

My research in the packet analyzer with a specialized search engine design was approved for the Augmentation Awards for Science and Engineering Research Training (ASSERT) program at the University of Texas at San Antonio. This program allowed additional research into this field of designing complex packet analyzers for gateways and/or firewalls with a specialized search engine. In today's LAN protocol analyzers, they are capable of negotiating a host of protocols and providing a multitude of monitoring/recording features at a significant price. Currently there are over 52 different protocol analyzers ranging from $1000 to $63,000 dollars. The analyzers also range from software (less expensive) to hardware (more expensive) solutions. Packet analyzers are used to help maintain networks, which include performance analysis of networks, network intrusion detection, network traffic logging, and fault analysis in networks.

The components of a packet analyzer are the hardware, capture driver, buffer, real-time analysis, decode and packet editing/transmission. Most products use standard hardware with the exception of packet analyzers that perform errors like CRC, negotiation, voltage, and cable problems. The capture driver is the most important part of the system. It captures the network traffic from the wire, filters it for the particular traffic and stores the data in a buffer. Once the frame is captured from the network, it is stored for further analysis. This real-time analysis sifts the traffic to find network performance issues and faults. The next component of analyzers is to show the packet information in a decoded format so the analyst can figure out what is going on. The last component is the ability to allow the analyzer to edit the network packets and transmit them onto the network.

PROBLEM

There is a strong need to capture and analyze network packets in today's home automation industry for several reasons. Several industry leaders are responding both with high dollar solutions to packet analyzers in the hardware realization. The problem with current packet analyzer implementations is that the analyzers are usually very difficult to configure, modify, and maintain. Home automation and control deals with "light-weight devices with very limited computing resources with high cost sensitivity. Control and automation networks also need to be more reliable and secure than the other forms of home networking. Given all the above, it can be seen that there is a need for architecture which addresses this situation. Current industry needs to develop a high-speed hardware packet filtering gateways that is cost effective and simple to maintain.

This research project addresses the design for architecture for this specialized gateway, which incorporates network packet analyzers and customized search engine for home automation data packets between computers on the same Ethernet circuit. When computers communicate over the network, all data is transmitted in small packets called frames. For Ethernet this maximum packet size is 1514 bytes where the Ethernet header is 14 bytes so this leaves 1500 bytes of data.

SCOPE

This project encompasses the design and implementation of a hardware realization of Ethernet packet analyzer with a customized search engine that is specific for the home automation industry. This analyzer will be at the gateway of a network and analyze Ethernet packets as they go by. It will keep a record of all packets sent to the gateway. It will also perform a real-time specialized search of the data packets for information that is for the home automation and not the computer network. This system will be a stand-alone real-time network analyzer system capable of decoding Ethernet protocol and interface via Cat 5 (RJ-45) cable.

The research involved in this project includes reviewing several different types of control architecture, clarifying search engine, protocols, and network interface controllers. The research encompassed literature searches from several universities and private industry. Also, interviewing several companies in the industry for current requirements of a design to meet future needs.

SUMMARY OF RESEARCH IN SEARCH ENGINES

The research into search engines included reviewing several documents and white papers on different ways of classification of documents and text. One of the documents call "KPS— A Web information Miningn Algorithm" addressed the semi-structured information. It states that it is not easy to search and extract structural data hidden in web pages or data packets. It all addressed the current practices that address the problem, which is syntax analysis or user-defined declaratives. In this paper, they present a novel information-mining algorithm, called KPS, which employs keywords, patterns, and/or samples to mine the desired information. They showed experimental results that KPS is more efficient than existing extracting methods.

Another document, "An Automatic Indexing and Neural Network Approach to Concept Retrival and Classification of Multilingual Document" addressed the automatic indexing and concept classification approach to a multilingual database. They introduced a multi-linear term-phrasing technique to extract concept descriptors (terms or keywords) from Chinese-English bibliographic database. This document discussed the problems with information retrieval and some of the techniques, such as, cluster analysis or vector space model; concept space clustering; and information space. Their research proposed a systematic blueprint of a multilingual classification model to help automatically index and classify unstructured information. Special features of this information classification model include a term phrasing and indexing framework components.

Another document, " Clarifying Search: A User-Interface Framework for Text Searches" addressed current user interfaces for textual database searching. The authors proposed a four-phase framework for user-interface design: (1) formulation, (2) action, (3) review of results, & (4) refinement. They also make recommendations as to how to implement the phases, based on the user's perspective; and show how two existing systems could be redesigned in accordance with their recommendations. They believe that designers armed with this information will be in a good position to satisfy the needs of all users.

Another paper, "An Implementation Study of a Detection-Based Adaptive Block Replacement Scheme" addressed a new adaptive buffer management scheme called DEAR (DEtection based Adaptive Replacement). This scheme automatically detects the block reference patterns of applications and applies different replacement policies. The authors implemented the DEAR scheme and measured its performance using several real applications. The results show that compared with the LRU buffer management scheme, the DEAR reduces the number of disk I/O.

These documents and other white papers gave some good insight into the research problem in this project. The final design of this project will include some of the ideas expressed by the authors.

SUMMARY OF NETWORK INTERFACE CONTROLLERS

The research into the network interface controllers included several different types of controllers like the DP83902 by National Semiconductor, LAN91C95 by Standard Microsystems Corporation, TDA5051A by Philips, and DP83934 by National Semiconductor.

The DP83902 is a serial network interface controller for twisted pair. This is a VLSI device designed for easy implementation of XSMA/CD local area networks. This device includes the receiver, transmitter, collision, heartbeat, loopback, jabber, and link integrity. This chip is a comprehensive single chip solution for networks and is designed for easy interface to other transceivers via the AUI interface.

The LAN91C95 is a VLSI Ethernet Controller that combines ISA and PCMCIA interfaces in one chip. This device integrates the entire MAC and the physical layer functions as well as the packet RAM needed to implement a high performance node. A unique architecture allows the LAN91C95 to combine high performance, flexibility, high integration and simple software interface. The Memory Management Unit (MMU) architecture combines the simplicity and low overhead of fixed areas with the flexibility of linked lists providing improved performance over other methods.

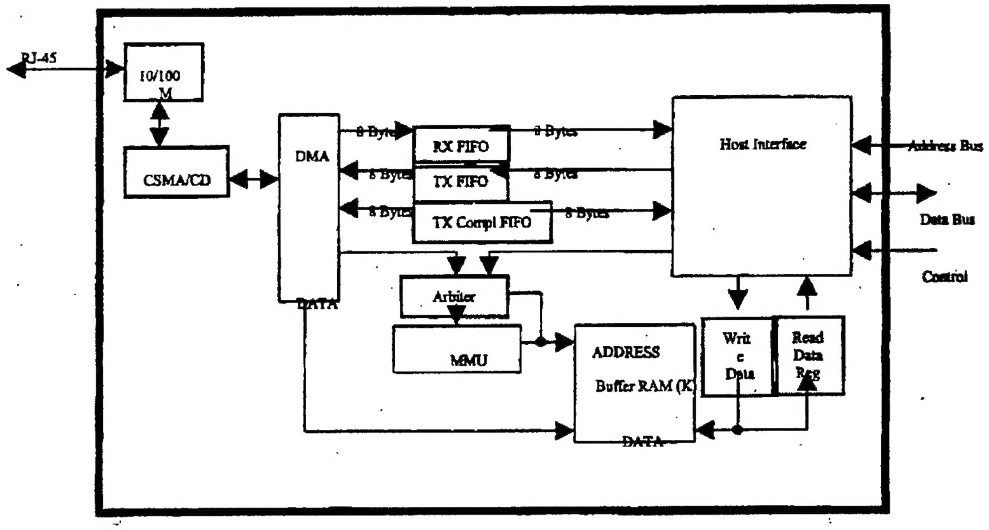
The TDA5051A is a modem IC, specifically dedicated to ASK transmission by means of the home power supply network. Both transmission and reception stages are controlled either by the master clock of the microcontroller, or by the on-chip reference oscillator connected to a crystal. The applications of this chip are for home appliance control, energy/heating control, and Amplitude Shift Keying (ASK) data transmission using the home power network.

The DP83934 is a SONIC-T second-generation Ethernet controller. This device is designed to meet high-speed 32- and 16-bit systems. Its system interfaces operates with high speed DMA that typically consumes less than 5% of the bandwidth. The linked-list buffer management system offers maximum flexibility in a variety of environments. For increased performance, the device implements a scheme to efficiently process receive and transmit packets in system memory. The receiver buffer management uses three areas in memory.

The above are only some of the devices that this research reviewed. The research revealed several different protocols for buffer memory management, which will be employed into my design. The network controllers also showed how to implement several modules of the network controllers that will be modified for my project.

SUMMARY OF ARCHITECTURE

The architecture for implementation uses data-flow model instead of a functional model. The data-flow model is also easier to partition and schedule in a real time limited resource implementation. The following is the architecture for this design:



The architecture will be futher explain and discussed in the final report or the master thesis for this project.

SUMMARY OR CONCLUSION

In summary, the research preformed by this project were in several areas: design for architecture for this specialized gateway, which incorporates network packet analyzers and customized search engine for home automation data packets between computers on the same Ethernet circuit.

The research for this project will directly relate to the completion of the master thesis and further pursuit of new ideas or solutions to the current industry requirements and problems. The research schedule for the work accomplished started in January 2000 to August 2000.

ADVISOR (S)

Dr. Parimal Patel, UTSA Professor, 210-458-5568
Dr Lin, UTSA Professor, 210-458-5528